Annu. Rev. Genomics Hum. Genet. 2002. 3:263–92
doi: 10.1146/annurev.genom.3.022502.103149
First published online as a Review in Advance on June 6, 2002

# BALANCED POLYMORPHISM SELECTED BY GENETIC VERSUS INFECTIOUS HUMAN DISEASE*

Michael Dean[1], Mary Carrington[2], and Stephen J. O'Brien[1]
*[1]Laboratory of Genomic Diversity and [2]Intramural Research Support Program, Science Applications International Corporation, National Cancer Institute, Frederick, Maryland 21702-1201; email: dean@ncifcrf.gov, carringt@ncifcrf.gov, obrien@ncifcrf.gov*

**Key Words** infection, HIV-1, HLA, chemokine receptors

■ **Abstract** The polymorphisms within the human genome include several functional variants that cause debilitating inherited diseases. An elevated frequency of some of these deleterious mutations can be explained by a beneficial effect that confers a selective advantage owing to disease resistance in carriers of such mutations during an infectious disease outbreak. We here review plausible examples of balanced functional polymorphisms and their roles in the defense against pathogens. The genome organization of the chemokine receptor and *HLA* gene clusters and their influence on the HIV/AIDS epidemic provides compelling evidence for the interaction of infectious and genetic diseases in recent human history.

## INTRODUCTION

Charles Darwin proposed that disease plays a role in the relative survival of species. However, Haldane proposed the concept that selection by infectious disease influences the frequency in genes that defend against pathogenic microbes (47). Although this concept is widely accepted today, there are few clear examples of infectious disease affecting allele frequencies, as this area of population genetics and evolution is particularly difficult to study in model systems. In addition, because the selective forces act over many generations and our knowledge of the history of human populations and infectious disease is incomplete over the last 10,000–100,000 years, these events are hard to reconstruct.

Natural selection for resistance to a pathogen can lead to the increase in frequency of alleles that are otherwise deleterious. An allele that reduces the risk of infection or pathogenesis by an infectious agent would generally involve a structural or regulatory alteration of a host protein. When that protein is involved in a vital function, then the resistant individual may acquire a deleterious phenotype. If

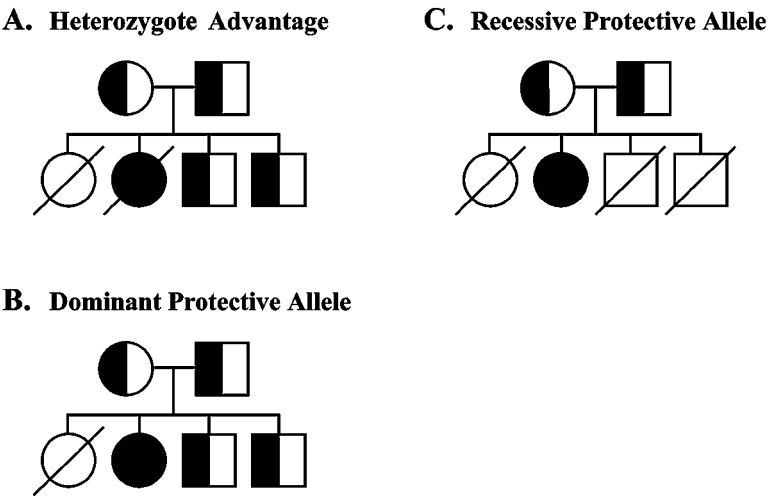



**Figure 1** Models for disease allele segregation. (*A*) Heterozygote advantage. An allele provides a selective advantage in the heterozygote but has a severe or lethal outcome in alternative homozygotes. In this situation, the frequency of heterozygotes gradually increases in the population. The *HbS* and other globin alleles protective effects for malaria are good examples. Circles, females; squares, males; half-filled symbols, carriers (heterozygotes); filled symbols, homozygotes; slashes, individuals who have died. (*B*) Dominant neutral allele. The allele providing the selective advantage for infectious disease is not deleterious. In this situation, homozygotes for the allele can accumulate and reach fixation in cases of a strong selective pressure. The *FY* null allele protective effect for malaria is an example of this model. (*C*) Recessive allele. In this situation, protection is only provided for homozygous individuals, and the variant is not deleterious to heterozygotes. *CCR5*-Δ*32* protection against AIDS infection is an example of this model.

protection from infection is a stronger selective force than the negatively selected phenotype, the deleterious allele will accumulate in the population as long as the infectious agent is epidemic. If heterozygous individuals are resistant to disease and homozygous individuals are severely compromised, heterozygote advantage is observed (Figure 1). If the infectious agent remains in the environment, selection to favor the deleterious allele continues.

Heterozygote advantage may play a role in the appearance of a large number of common severe genetic diseases (sickle cell anemia, cystic fibrosis, Tay-Sachs disease, and others). Recessive diseases fit the heterozygote advantage paradigm best. However, owing to historical fluctuations in the effective size of human populations and the known propensity for restricted mating, recessive deleterious alleles may reach elevated frequencies as a result of genetic drift (141). Recessive alleles manifest a burden on the population only when they reach a high enough

frequency to result in the appearance of homozygous individuals. For example, at an allele frequency of 10%, a recessive lethal allele results in only a 1% increase in mortality. In the absence of clear evidence for selective maintenance of rare alleles, making a distinction between selective influence and genetic drift is difficult.

## Malaria Resistance Genes

The selection for resistance to malaria has been extensively studied, and overwhelming evidence that Plasmodium has shaped the human genome exists. The best-studied allele is the *HbS* allele of the $\beta$-globin gene (60). This allele is responsible for sickle cell disease in heterozygotes and sickle cell anemia in homozygotes. Sickle cell anemia is a very serious disease, resulting in characteristically misshaped red blood cells (51). The *Hbs* allele is quite common in populations of African origin, particularly in regions where malaria is endemic. Sickle cell disease anemia results in several rare, mostly late onset, pathologies, including stroke, sudden unexplained death, and acute chest syndrome (1). Therefore, practically all carriers survive to reproductive age, and there is no significant reduction in fertility. Heterozygotes often give birth to newborns affected with anemia. Although current treatments can extend the lifespan of affected individuals, in the absence of therapy, the disease is lethal. In addition to sickle cell disease, $\alpha$- and $\beta$-thalassemia alleles are common in malaria endemic regions and are also thought to arise as a result of malaria selection (34).

The HbS protein was one of the first human disease proteins characterized. Research on the gene also provided early examples of disease allele identification, genetic testing (66), and linkage disequilibrium analysis. In 1949, Haldane proposed that the allele might have arisen in response to a selective agent, and he proposed malaria as that agent (47). Population data supports this finding, as do biochemical studies on the interaction of the falciparum malaria parasite with red blood cells (35). Thus, HbS remains the best-known example of a homozygous lethal allele providing a heterozygote advantage.

The *FY*/*Duffy* gene encodes a chemokine receptor that expresses a red blood cell antigen that is polymorphic and found at widely different frequencies in world populations. A null allele of *FY* occurs at up to 100% frequency in African populations, whereas it is rare to nonexistent in Asian and Caucasian populations. The FY protein is a receptor for vivax malaria, and therefore, homozygous null individuals are resistant to infection (94). No deleterious phenotype has been ascribed to FY heterozygotes or homozygotes (22, 46). The *FY* allele provides an example of a selectively neutral allele whose frequency was elevated by its conferring resistance to an infectious agent.

## Cystic Fibrosis and Other Diseases

Cystic fibrosis (CF) is a fatal monogenic recessive disease, reaching frequencies as high as 1/1600 in Caucasian populations, corresponding to a heterozygote carrier frequency of 1/20. The disease is much less common in African, Arabic, and

Asian populations, and carrier frequencies of 1/100 to 1/200 have been calculated (32). Two pieces of genetic evidence would suggest a selective advantage for CF alleles: the geographical/population differences in the frequency of the disease, and the appearance of distinct alleles at high frequency in isolated populations. The frequency of CF is highest in Northern Europe and declines in both the South and East (137). This corresponds with the frequency of the major allele of the responsible *CFTR* gene, a deletion of three base pairs that removes the phenylalanine at position 508 (ΔF508) (69). The ΔF508 allele accounts for 70%–80% of CF alleles in Britain but only 40%–50% in Greece and Russia (32). However, at least two other populations have high frequency CF alleles. The W1282X allele is found at a frequency of 51% in the Ashkenazim (65). The 1677delTA allele has been found at a high frequency in Georgians and is also present at an elevated level in Turkish and Bulgarian populations (3). Because genetic drift might contribute to allele frequency elevations, even for a lethal allele, gene frequency differences alone cannot be taken as proof for a heterozygote advantage of CF alleles.

The *CFTR* gene encodes for a chloride ion channel that plays a major role in the regulation of exocrine secretions. Through the CFTR channel, bacterial toxins, such as cholera and *Escherichia coli*, cause increased fluid flow in the intestine and result in diarrhea. Several researchers have proposed that the CF mutations have been selected in response to these diseases. CF homozygotes fail to secrete chloride ions in response to a variety of stimulants, including bacterial toxins, and mice heterozygous for *CFTR* null alleles show reduced intestinal fluid secretion in response to cholera toxin (36), lending support to the infectious disease influence on CFTR allele frequencies. However, a subsequent study of *CFTR* heterozygous mice did not reveal a defect in intestinal secretion (25), and *CFTR* expression does not appear to be a rate-limiting step in chloride secretion in humans (54). An alternative hypothesis is that CFTR is involved in resistance to typhoid fever. *Salmonella typhi*, but not the related *S. typhimurium*, uses CFTR to enter epithelial cells (111). In one experiment, mice heterozygous for *Cftr-ΔF508* showed an 86% reduction in internalization of *S. typhi* into the gastrointestinal tract, relative to that in wild-type mice, and antibodies to CFTR inhibited uptake (111). Thus, selection for typhoid fever resistance provides a plausible explanation for the high frequency of *CFTR* mutations.

Ashkenazi Jewish populations have a very high frequency of several rare lysosomal storage disorders, including Tay-Sachs, Gaucher, and Neimann-Pick diseases. In each of these diseases, individual disease alleles have reached high frequency. For example, the 1277ins4 allele in *HEXA* represents 73%–79% of the Ashkenazi Tay-Sachs alleles (44). Several pieces of evidence point to a selective event(s) as the source of the high frequency of these alleles. Reduced incidence of tuberculosis among grandparents of Tay-Sachs homozygotes has been interpreted as suggestive that these higher deleterious alleles may be involved in resistance to TB (29, 102).

Selection is difficult to document in humans, as experimental intervention is limited, environmental conditions have changed dramatically in the last centuries, and many of the pathogens that severely affected human populations have been

eradicated. Also, even weak selective influences would result in a significant increase in allele frequency over many generations. Therefore, if infectious disease provides an important selective force and is responsible for the fixation of many deleterious alleles in the population, better understanding this process would seem critical.

## HIV/AIDS Restriction Genes

Human immunodeficiency virus (HIV-1), the cause of AIDS, has many advantages for the study of host genes that alter viral infection and disease (105). The AIDS pandemic is recent, widespread, and has affected several different risk groups (male homosexuals, intravenous drug users, hemophiliacs, commercial sex workers, and transfusion recipients). Studies of high-risk individuals consistently show that a subset fails to become infected, suggesting that there are some individuals with a lower risk of infection. In addition, the identification of individuals infected during their participation in a longitudinal epidemiological study has allowed the identification of acute seroconverters, which have precise dates of HIV-1 exposure and infection. This group can be studied by survival and relative hazard analyses that are highly sensitive in identifying genes that have modest effects on the rate of progression to full-blown AIDS (105). Finally, the intense research efforts on the biology of HIV-1 have identified several host proteins that are essential for viral infection and disease progression, providing plausible biological candidates for AIDS restriction genes.

The identification of chemokines as inhibitors of HIV-1 infection and the discovery that certain chemokine receptors are essential for viral infection has led to the intense study of these two classes of proteins (40). Chemokines comprise 80–100 amino acid proteins that are often secreted by damaged or infected cells (114). Receptors on the surface of lymphoid cells function to direct these cells to areas of infection and/or inflammation. A subset of chemokine receptors interact with the CD4 protein to form a complex that is recognized by the envelope protein of HIV-1 and allows the virus to fuse with the cell of the viral genome (10). Chemokines can compete for chemokine receptor binding and inhibit HIV-1 infection (18).

The primary chemokine receptors for HIV-1 are CCR5 and CXCR4 (10). Viruses that exclusively use CCR5 are known as R5 strains, and those that use CXCR4 are X4 strains. For reasons that are poorly understood, R5 strains are predominant in the initial infection of most individuals. As the infection progresses, X4 variants arise de novo by acquisition of new mutations in the HIV-1 envelope gene. The derivative HIV strain can spread to additional cell types, and the R5-X4 transition typically is a prelude to the rapid decline in immune function and to the appearance of AIDS.

Several groups identified a 32-bp deletion mutation in the coding region of the human *CCR5* gene that results in an inactive protein (26, 77, 119, 143). This *CCR5*-Δ*32* allele is found at frequencies of 5%–15% in Caucasian populations. Homozygotes are highly resistant to HIV-1 infection, but the protection from HIV-1

infection is not complete, as several infected *CCR5-Δ32* homozygous subjects have been described (104). *CCR5-Δ32* heterozygotes do not display a reduced risk for infection but do demonstrate a slower rate of disease progression after infection, when measured as a decline in CD4 positive T cells below 200/ml, the appearance of AIDS-defining conditions or death due to HIV-1 infection (26). Variants in the *CCR5* promoter are also associated with altered rates of AIDS progression, suggesting that the level of CCR5 is rate-limiting for viral spread (87).

The *CCR2* gene (21) is adjacent to *CCR5* on human chromosome 3, and certain cell culture–adapted strains of HIV-1 can use CCR2 to enter cells (23, 31). An allele of *CCR2* (*CCR2-64I*) was identified that is also associated with slower progression to AIDS (126). Because CCR2 does not play a major role in HIV-1 infection in vivo, the role of the *CCR2-64I* allele is unclear. However, the CCR2-64I form of the protein may differentially interact with CXCR4 and potentially act indirectly to affect HIV-1 replication (92).

The chemokine SDF-1 is the principal ligand for CXCR4 and can inhibit infection of X4 tropic strains of HIV-1 (101). A variant in the 3′ untranslated region of the *SDF1* gene (*SDF1-3′A*) that is associated with a delayed progression to AIDS was identified (140). The *SDF1-3′A* allele may be associated with altered SDF-1 protein levels, but this has not been directly demonstrated. The interleukin-10 (IL10) protein plays an inhibitory role in replication and proliferation of macrophages, T cells, and HIV-1. A variant (*IL10-5′A*) in the *IL10* promoter region that diminishes IL10 transcription is associated with accelerated progression to AIDS (122). Gel-shift experiments indicate that the *IL10-5′A* allele displays reduced binding to a nuclear protein (122). In total, six variants in chemokine, chemokine receptor, and cytokine genes (*CCR5*, *CCR5P*, *CCR2*, *SDF1*, *RANTES*, and *IL10*) account for a significant fraction of the heterogeneity in AIDS progression (90, 104, 105).

## GENOME ORGANIZATION OF THE CHEMOKINE GENES

The chemokine receptors are members of the seven-transmembrane G-protein-coupled receptor (GPCR) superfamily. Like all GPCRs, the chemokine receptors have an extracellular binding domain, a set of seven-transmembrance domains and an intracellular signaling domain (79). The CC-chemokine receptor genes typically have their coding region contained in one exon and often have one or more 5′ noncoding exons. The *CCR2* gene has differential use of a second exon at the 3′ end. The promoters for most of the genes have not been well characterized. In the case of the *CCR5* gene, there are several potential upstream noncoding exons and transcription start sites (78, 97).

At least 12 chemokine receptor genes are present in a cluster that spans over 8 Mb from 3p22.3-3p21.31 (85) (Figure 2). The *CCR5* and *CCR2* genes are separated by only 19 kb and the *CCR1* and *CCR3* genes by ∼50 kb. The region is moderately well covered by the genome sequencing projects, with only

the *CCR4* gene absent. The "private" Celera draft genome sequence has the entire region represented on a single 50.7-Mb scaffold (data not shown). All genes in the cluster have known ligands, except for CCBP2, a CCR10-related gene. The genes are of diverse origin and include both CC and CXC receptors. Most genes are facing in the same transcriptional orientation with their promoter telomeric to the 3′ end. Interspersed in this cluster of chemokine receptors are many other genes that are unrelated in both sequence and function. It is interesting to note that the *TYMSTR* gene is in the intron of the *FYC01* gene, a coiled-coil domain–containing protein. Thus, multiple gene duplication and rearrangement events that have led to the generation of this locus have occurred. Many of these genes' homologues are conserved as a cluster on mouse chromosome 9 (14).

Phylogenetic analysis of the chemokine receptor genes indicates a high degree of conservation between genes. The *CCR2* and *CCR5* genes are the most closely related, and the *CCR1* and *CCR3* genes also form a closely related group (142). Interestingly, *CCR2* and *CCR5* bind to nonoverlapping sets of ligands, so they have diverged functionally while retaining considerable structural similarity. This is consistent with these two pairs of genes having been derived from a recent duplication event. Sequence analysis of the chemokine receptor genes from other primate and mammalian species demonstrates that the genes have been highly conserved, consistent with the hypothesis that they encode essential functions (142) (W. Johnson, M. Dean, & S.J. O'Brien, manuscript in preparation).

## CHEMOKINE RECEPTOR POLYMORPHISMS

Many of the chemokine receptor genes have been examined extensively for genetic variation or single nucleotide polymorphism (SNP) (4, 19) (Table 1). In addition to the *CCR5*-Δ*32* allele, at least 20 other *CCR5* variants in the coding region of the gene are known (Table 2). Several of these variants are common, and most result in amino acid replacements. The L55Q allele replaces a conserved leucine residue for glutamine in the first transmembrane domain. This alteration results in a CCR5 protein that has a higher affinity for the ligands RANTES and MIP1$\alpha$ (56) and shows an increased frequency in HIV-1 infected subjects (M. Dean, unpublished observation). A rare nonsense mutation, C101X, in European populations (19, 113) and two frame-shift mutations in Asian populations (4) have been described. Replacement of a conserved cystine that forms a critical disulfide bond (C20S) was identified in several patients, including an HIV-1 negative, high-risk individual who was also heterozygous for *CCR5*-Δ*32* (19). In vitro data confirms that the C20S is defective in both chemokine binding and HIV-1 entry (13, 56). A surprising number of the nonsynonymous variants in *CCR5* alter the function of the protein for either HIV infection or chemokine binding and response (56).

The *CCR5* promoter has also been extensively surveyed for variation, and a large number of variable sites exist (45, 74, 87, 91, 98). There is extensive linkage disequilibrium (LD) between these variants that extends at least 20 kb through the *CCR2* gene and its promoter. These LD relationships are present in all major racial

**TABLE 1** Chemokine receptor genes: location and polymorphism

| Gene | Alias | dbSNP | Validation | Bp change | AA change | Cyto. | Genome location[a] |
|---|---|---|---|---|---|---|---|
| CCR4 | | none | | | | 3p24 | ? |
| CX3CR1 | | rs1050592 | N | C/T | | 3p22.3 | 48, 242, 307 |
| | | a | Y | G/A | V249I[b] | | |
| | | a | Y | C/T | T280M[b] | | |
| CCR8 | | none | | | | 3p22.3 | 48, 500, 433 |
| CCBP2 | CCR9, | rs1054138 | N | A/T | Q/L | 3p21.33 | 52, 488, 537 |
| | CCR10 | rs1054139 | N | A/T | Q/H | | |
| CCR9 | | rs1488371 | N | A/C | Intron | 3p21.32 | 55, 642, 305 |
| TYMSTR | STRL33 | rs1386931 | N | C/T | 3′ UTR | 3p21.32 | 55, 690, 834 |
| | | rs936939 | N | A/C | 3′ UTR | | |
| CCXCR1 | | rs876668 | N | A/G | promoter | 3p21.32 | 55, 793, 068 |
| CCR1 | | none | | | | 3p21.32 | 55, 944, 285 |
| CCR3 | | b | Y | T/C | Y17Y[c] | 3p21.31 | 56, 038, 921 |
| CCR2 | | rs1799864 | Y | G/A | V64I | 3p21.31 | 56, 138, 349 |
| | | c | Y | C/T | N260N[d] | | |
| | | c | Y | C/T | P47L[d] | | |
| | | c | Y | G/T | V52V[d] | | |
| | | c | Y | T/G | S87A[d] | | |
| | | rs762789 | N | A/G | 3′ flanking | | |
| | | rs1042100 | N | A/G | 3′ UTR | | |
| | | rs743660 | N | A/G | 3′ UTR | | |
| | | rs1042108 | N | G/A | 3′ UTR | | |
| CCR5 | | see Table 2 | | | | 3p21.31 | 56, 151, 580 |
| CCRL2 | CCR6, CCRX | None | | | | 3p21.31 | 56, 188, 736 |

[a]Location in Santa Cruz Assembly (http://genome.ucsc.edu/) in bp from beginning of the chromosome. CCR4 is not represented in the genome sequence.

[b]Described in (67a).

[c]Described in (33a).

[d]Described in (126). Validation polymorphism is described in large-scale SNP detection/prediction efforts (N) or has been independently documented (Y).

groups, suggesting that very little recombination occurs in this genetic interval. In fact, the *CCR2* and *CCR5* genes lie at the boundary of a region that extends 5 Mb telomeric to the cluster that displays very low recombination (Figure 3).

The *CCR2-64I* allele is in nearly complete association with the 59653T variant in the *CCR5* promoter region. Common variants are also found at positions 58755, 58934, 59029, 59353, and 59356. Four haplotypes account for over 90% of the haplotypes observed (87). Although there is significant difference between large racial groups in the frequencies of these major haplotypes, the same haplotypes are present in all of the populations. The finding that varying haplotypes of the CCR5

**TABLE 2** Genetic variants of the *CCR5* gene*

| Variant | Nucleic acid substitution | Frequency African American | Frequency Caucasian | References |
|---|---|---|---|---|
| Promoter | 58755 A/G | 0.06 | 0.11 | (87) |
| Promoter | 58934 G/T | 0.67 | 0.65 | (87) |
| Promoter | 59029 A/G | 0.44 | 0.56 | (87) |
| Promoter | 59029 G/T | <0.01 | <0.01 | (87) |
| Promoter | 59338 A/G | <0.01 | <0.01 | (87) |
| Promoter | 59352 C/A | <0.01 | <0.01 | (87) |
| Promoter | 59353 C/T | 0.44 | 0.56 | (87) |
| Promoter | 59356 C/T | 0.67 | 0.65 | (87) |
| Promoter | 59373 C/A | <0.01 | <0.01 | (87) |
| Promoter | 59402 A/G | 0.15 | 0.35 | (87) |
| Promoter | 59410 T/C | <0.01 | <0.01 | (87) |
| Promoter | 59440 C/G | <0.01 | <0.01 | (87) |
| Promoter | 59537 G/A | <0.01 | <0.01 | (87) |
| Promoter | 59653 C/T | 0.16 | 0.10 | (87) |
| I12T | 25 A/C | <0.01 | 0.003 | (86) |
| C20S | 58 T/A | <0.01 | 0.003 | (86) |
| A29S | 85 G/T | 0.01 | <0.01 | (86) |
| I42F | 24 A/T | <0.01 | 0.001 | (86) |
| L55Q | 164 T/A | 0.007 | 0.04 | (85, 86) |
| R60S | 180 G/T | 0.01 | <0.01 | (86) |
| S63C | 187 A/T | <0.01 | <0.01 | (87) |
| A73V | 218 C/T | <0.01 | 0.002 | (86) |
| S75S | 215 T/C | 0.013 | <0.01 | (86) |
| C101X | 303 T/A | <0.01 | 0.01 | (86) |
| Δ32 (185) | Δ32 | 0.02 | 0.08 | |
| L215S | 664 C/T | <0.01 | <0.01 | (85) |
| R223Q | 668 G/A | <0.01 | 0.01 | (85, 86) |
| 228delK | 680del3 | <0.01 | 0.002 | (86) |
| 299FS | 893delC | <0.01 | <0.01 | (85) |
| V300V | 900 C/A | 0.004 | <0.01 | (86) |
| G301V | 902 G/T | <0.01 | 0.01 | (86) |
| R319H | 956 G/A | <0.01 | <0.01 | (87) |
| P332P | 996 C/T | <0.01 | <0.01 | (85) |
| A335V | 1004 C/T | 0.025 | 0.006 | (85, 86) |
| Y339F | 1016 A/T | 0.03 | <0.01 | (85, 86) |

*The promoter and 5′ UTR variant are numbers according to BAC clone U95626. The coding region variants are numbered with bp 1 as the first base of the CCR5 initiation codon. Variants in the CCR5 promoter (87, 91, 97) and CCR5 coding region (4, 19, 113) have been described previously.

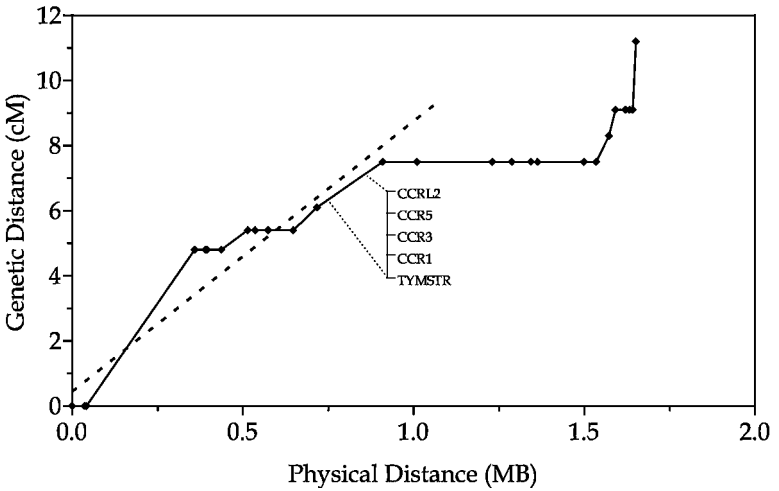



**Figure 3** Genetic versus physical distance in the CCR region. A plot of genetic distance in centimorgans (cM), and physical distance in bp is shown for microsatellite markers in the CCR region. The genetic distances are from the GeneMap99 (http://www.ncbi.nlm.nih.gov/genemap99) and the physical distances from the Human Genome Project Working Draft. The map spans from D3S1260 to D3S2408. The location of several chemokine genes is shown. The straight line represents 1 cM/Mb.

promoter region are associated with differential progression to AIDS suggests that these haplotypes are not functionally similar. This is further supported by gel-shift experiments that show differential binding of nuclear proteins at some of these polymorphic sites (17).

The *CCR2* and other chemokine receptor genes have been less intensively interrogated; however besides *CCR2-64I*, one common synonymous variant, as well as several rare nonsynonymous variants, is present in the coding region (Table 1). The *CCR2* promoter region contains many SNPs that are also common (V. Clark, M. Dean, unpublished results). Several of the other chemokine receptor genes have at least one common nonsynonymous allele. Several of these alleles, including *CCR2-64I* and *BOB/GPF15* genes, display altered function (Table 1).

## POPULATION DISTRIBUTION OF *CCR2* AND *CCR5* VARIANTS

The *CCR5-Δ32* allele is found almost exclusively in populations of European origin as well as populations that have undergone admixture with Caucasians (Table 3). In addition, the allele is found predominantly on a single haplotype, consistent with the notion that it arose once in the population, after Caucasians

**TABLE 3** *CCR5-Δ32* frequency in worldwide populations

| Population | +/+ | +/Δ | Δ/Δ | Sum | Freq. Δ-*32* |
|---|---|---|---|---|---|
| Europe | | | | | |
| Mordvinian | 58 | 28 | 0 | 86 | 0.16 |
| Iceland | 75 | 24 | 3 | 102 | 0.15 |
| Sweden | 251 | 74 | 10 | 335 | 0.14 |
| Slovakia | 22 | 8 | 0 | 30 | 0.13 |
| Estonia | 116 | 42 | 0 | 158 | 0.13 |
| Russian | 141 | 43 | 2 | 186 | 0.13 |
| Ashkenazi | 721 | 209 | 19 | 949 | 0.13 |
| Denmark | 387 | 104 | 7 | 498 | 0.12 |
| Britain | 553 | 142 | 10 | 705 | 0.11 |
| Lithuania | 220 | 61 | 2 | 283 | 0.11 |
| Poland | 694 | 190 | 7 | 891 | 0.11 |
| Finland | 228 | 65 | 0 | 293 | 0.11 |
| West Siberian | 86 | 13 | 5 | 104 | 0.11 |
| France-Brittany | 79 | 20 | 1 | 100 | 0.11 |
| Germany | 168 | 45 | 1 | 214 | 0.11 |
| France-North | 1044 | 229 | 26 | 1299 | 0.11 |
| Czech | 434 | 109 | 4 | 547 | 0.11 |
| Norway | 79 | 21 | 0 | 100 | 0.11 |
| US Caucasian | 2677 | 460 | 26 | 2496 | 0.10 |
| Netherlands | 291 | 73 | 0 | 364 | 0.10 |
| Spain | 45 | 11 | 0 | 56 | 0.10 |
| Hungary | 179 | 34 | 4 | 217 | 0.10 |
| Spain-Murcia | 81 | 19 | 0 | 100 | 0.10 |
| South African Caucasian | 118 | 25 | 1 | 144 | 0.09 |
| Belgium | 582 | 114 | 8 | 704 | 0.09 |
| Switzerland | 54 | 9 | 1 | 64 | 0.09 |
| Austria | 30 | 6 | 0 | 36 | 0.08 |
| Sweden-Saami | 101 | 18 | 1 | 120 | 0.08 |
| Slovenia | 465 | 88 | 2 | 555 | 0.08 |
| Albania | 62 | 10 | 1 | 73 | 0.08 |
| Spain-Catalan | 115 | 22 | 0 | 137 | 0.08 |
| France-South | 655 | 105 | 6 | 766 | 0.08 |
| Italy-North | 137 | 22 | 1 | 160 | 0.08 |
| Ireland | 64 | 11 | 0 | 75 | 0.07 |
| Spain-Basque | 103 | 14 | 1 | 118 | 0.07 |
| Portugal | 88 | 13 | 0 | 101 | 0.06 |
| Caucasus-Daghestan | 96 | 14 | 0 | 110 | 0.06 |
| Italy | 234 | 29 | 0 | 263 | 0.06 |
| Bulgaria | 40 | 4 | 0 | 44 | 0.05 |
| Romania | 10 | 1 | 0 | 11 | 0.05 |
| Sardinia | 92 | 8 | 0 | 100 | 0.04 |
| Greece | 537 | 37 | 1 | 575 | 0.03 |

(*Continued*)

**TABLE 3** (*Continued*)

| Population | +/+ | +/Δ | Δ/Δ | Sum | Freq. Δ-*32* |
|---|---|---|---|---|---|
| Bulgaria-Gypsy | 44 | 3 | 0 | 47 | 0.03 |
| Cyprus | 364 | 23 | 0 | 387 | 0.03 |
| Georgia | 50 | 0 | 0 | 50 | 0.00 |
| Middle East | | | | | |
| Turkey | 126 | 18 | 0 | 144 | 0.06 |
| Israel-Sephardic | 675 | 66 | 10 | 751 | 0.06 |
| Saudia Arabia | 331 | 10 | 0 | 341 | 0.01 |
| Yemen | 123 | 0 | 1 | 124 | 0.01 |
| Lebanon | 51 | 0 | 0 | 51 | 0.00 |
| Asia | | | | | |
| Russia-Tatar | 38 | 12 | 0 | 50 | 0.12 |
| Russia-Udmurtia | 38 | 7 | 1 | 46 | 0.10 |
| Azerbaijan | 36 | 4 | 0 | 40 | 0.05 |
| Gujurat | 30 | 1 | 1 | 32 | 0.05 |
| Uzbekistan | 27 | 2 | 0 | 29 | 0.03 |
| Asian-American | 62 | 4 | 0 | 66 | 0.03 |
| Kazakistan | 47 | 3 | 0 | 50 | 0.03 |
| Pakistan | 32 | 2 | 0 | 34 | 0.03 |
| Russia-Uyguryi | 43 | 2 | 0 | 45 | 0.02 |
| Russia-Tuvinian | 48 | 2 | 0 | 50 | 0.02 |
| Sindh | 28 | 1 | 0 | 29 | 0.02 |
| Asian Jews | 62 | 3 | 0 | 65 | 0.02 |
| Punjab | 33 | 1 | 0 | 34 | 0.01 |
| East Siberian | 214 | 5 | 0 | 219 | 0.01 |
| India | 99 | 1 | 0 | 100 | 0.01 |
| Central Asian | 106 | 1 | 0 | 107 | 0.00 |
| Thailand | 1154 | 7 | 0 | 1161 | 0.00 |
| China | 446 | 1 | 0 | 447 | 0.00 |
| Bengal | 25 | 0 | 0 | 25 | 0.00 |
| Taiwan | 425 | 0 | 0 | 425 | 0.00 |
| Phillipines-Filipino | 26 | 0 | 0 | 26 | 0.00 |
| Philipines-Nigrino | 30 | 0 | 0 | 30 | 0.00 |
| Mongolia | 59 | 0 | 0 | 59 | 0.00 |
| Sri Lanka | 37 | 0 | 0 | 37 | 0.00 |
| Burma | 67 | 0 | 0 | 67 | 0.00 |
| Borneo | 151 | 0 | 0 | 151 | 0.00 |
| Sumatra | 72 | 0 | 0 | 72 | 0.00 |
| Andaman Islands | 24 | 0 | 0 | 24 | 0.00 |
| Kota-Kinabalu | 178 | 0 | 0 | 178 | 0.00 |
| Korea | 504 | 0 | 0 | 504 | 0.00 |
| Japan | 248 | 0 | 0 | 248 | 0.00 |
| Africa | | | | | |
| African-American | 311 | 10 | 0 | 321 | 0.02 |

(*Continued*)

**TABLE 3** (*Continued*)

| Population | +/+ | +/Δ | Δ/Δ | Sum | Freq. Δ-*32* |
|---|---|---|---|---|---|
| Morocco | 163 | 3 | 1 | 167 | 0.01 |
| West Africa | 137 | 2 | 0 | 139 | 0.01 |
| Nigeria | 110 | 1 | 0 | 111 | 0.00 |
| Gambia | 103 | 0 | 0 | 103 | 0.00 |
| Central African Republic | 52 | 0 | 0 | 52 | 0.00 |
| Kenya | 87 | 0 | 0 | 87 | 0.00 |
| Tanzania | 7 | 0 | 0 | 7 | 0.00 |
| Ivory Coast | 87 | 0 | 0 | 87 | 0.00 |
| Malawi | 93 | 0 | 0 | 93 | 0.00 |
| Zambia | 96 | 0 | 0 | 96 | 0.00 |
| Kalari San | 36 | 0 | 0 | 36 | 0.00 |
| Uganda | 6 | 0 | 0 | 6 | 0.00 |
| Madagascar-Merina | 42 | 0 | 0 | 42 | 0.00 |
| Madagascar-Bezanozano | 16 | 0 | 0 | 16 | 0.00 |
| Madagascar-Betsileo | 42 | 0 | 0 | 42 | 0.00 |
| Madagascar-Sihanaka | 19 | 0 | 0 | 19 | 0.00 |
| Africa | 150 | 0 | 0 | 150 | 0.00 |
| Africa-West, Central | 124 | 0 | 0 | 124 | 0.00 |
| Oceania | | | | | |
| Pohnpei | 28 | 1 | 0 | 29 | 0.02 |
| Australian Aboriginals | 96 | 2 | 0 | 98 | 0.01 |
| Pacific Island | 481 | 9 | 0 | 490 | 0.01 |
| Guam | 58 | 1 | 0 | 59 | 0.01 |
| New Guinea Coast | 96 | 0 | 0 | 96 | 0.00 |
| French Polynesian | 94 | 0 | 0 | 94 | 0.00 |
| Fiji | 17 | 0 | 0 | 17 | 0.00 |
| Banks and Torres Islands | 38 | 0 | 0 | 38 | 0.00 |
| Vanuatu-Maewo | 22 | 0 | 0 | 22 | 0.00 |
| Vanuatu-Santo | 18 | 0 | 0 | 18 | 0.00 |
| Nauru | 30 | 0 | 0 | 30 | 0.00 |
| Majuro | 29 | 0 | 0 | 29 | 0.00 |
| Truk | 30 | 0 | 0 | 30 | 0.00 |
| Palau | 30 | 0 | 0 | 30 | 0.00 |
| Kiribati | 30 | 0 | 0 | 30 | 0.00 |
| Kapingamarangi | 30 | 0 | 0 | 30 | 0.00 |
| Tonga | 36 | 0 | 0 | 36 | 0.00 |
| Americas | | | | | |
| Mexico | 134 | 11 | 0 | 145 | 0.04 |
| Brazil | 93 | 7 | 0 | 100 | 0.04 |
| Puerto Rico | 286 | 17 | 0 | 303 | 0.03 |
| Columbia | 207 | 10 | 1 | 218 | 0.03 |
| Hispanic American | 354 | 20 | 0 | 374 | 0.03 |
| Haiti | 82 | 3 | 0 | 85 | 0.02 |

(*Continued*)

**TABLE 3** (*Continued*)

| Population | +/+ | +/Δ | Δ/Δ | Sum | Freq. Δ-*32* |
|---|---|---|---|---|---|
| Nuu-Chah-Nulth | 37 | 1 | 0 | 38 | 0.01 |
| Mexico-Huicholes | 52 | 0 | 0 | 52 | 0.00 |
| Brazilian Amerindians | 398 | 0 | 0 | 398 | 0.00 |
| Jamaica | 119 | 0 | 0 | 119 | 0.00 |
| Cheyenne | 20 | 0 | 0 | 20 | 0.00 |
| Pima | 40 | 0 | 0 | 40 | 0.00 |
| Pueblo | 20 | 0 | 0 | 20 | 0.00 |
| Inuit | 40 | 0 | 0 | 40 | 0.00 |
| Total | 23,091 | 2851 | 165 | 26107 | 0.06 |

Data derive principally from large surveys of populations (76, 80, 81, 88, 89, 127). A complete table with all references is available upon request.

diverged from other racial groups (76, 127). Within European populations, there is a significant geographic gene frequency variance in allele frequency, with the highest frequencies (12%–14%) found in Northern Europe and the lowest (4%–6%) in the South (76, 89, 127). Microsatellite markers and coalescence analysis have been used to estimate the age of currently observed *CCR5*-Δ*32*-bearing haplotypes. The results indicate that the allele is of recent origin, perhaps as recent as 800 years ago (76, 127). Given the high frequency and widespread distribution of this allele, this possibility is consistent with strong selective pressure favoring an increase in *CCR5*-Δ*32* frequency. The recombination distance between *CCR5* and adjacent microsatellite markers was imputed from a regression radiation hybrid physical distance and recombination distances (127). From the current genome data, these physical distances can be refined (Figure 2) and produce an estimate of the age of *CCR5*-Δ*32*-bearing haplotype as 1400 years. This estimate, however, remains tentative as it is based on an imprecise relationship of physical and recombination distance in a region of known recombination reduction (Figure 2). As such, a better estimate based on actual recombination (e.g., from sperm typing) (41) would be useful to refine the age of this haplotype.

The Ashkenazi Jewish (73a) and Tatar populations also have a high prevalence of the *CCR5*-Δ*32* allele (127) (Table 3). These ethnic groups reside in the North of Europe, adjacent to European populations with high *CCR5*-Δ*32* allele frequencies. The Tatars, native to Siberia, are of Mongolian origin, whereas the Ashkenazim are believed to have separated from the rest of the Caucasian population approximately 2000 years ago. It is tempting to speculate that a combination of admixture and a common selective force led to the increased frequency of this allele in these populations.

Because CCR5 is the ligand for several immune response signaling molecules, an infectious agent seems more likely to be the selective agent rather than a noninfectious force, such as fertility or fetal development. Although HIV-1 is believed

to have arisen in Africa in the early twentieth century (37, 117), a related lentivirus may have been present in Europe but did not survive to the present. Lenitviruses are prevalent in most primates and other mammals, and several cross-species transmissions have been documented. Other potential selective agents include a wide range of infectious diseases, such as smallpox, tuberculosis, yellow fever, rubella, typhus, malaria, cholera, polio, and bubonic plague, which have affected European populations. Two pieces of information implicate plague as a potential source of a CCR5 mutation selection. The infectious agent that causes plague, *Y. pestis*, infects macrophages, the principal location of CCR5; and the European locale and timing of the major plague epidemics coincide with the calculated origin of the *CCR5-Δ32* allele. Animal experiments with a *Ccr5−/−* mouse and/or in vitro tests of human cells with alternative *CCR5* genotypes will be required to validate this hypothesis.

## CHEMOKINE RECEPTOR GENE AND HUMAN DISEASE

The presence of the cluster of chemokine receptor genes on chromosome 3p21-24 and the high frequency of functional alleles in these genes suggest that these genes might be good candidates as modifying or susceptibility alleles for complex diseases. Several mendelian disorders map to this region (Table 4). However, most involve deafness (DFNB6), skeletal development (ARRVD5), or encephalopathy (AS1); phenotypes not obviously linked to chemokine action. This region of chromosome 3p is subject to loss-of-heterozygosity in a large portion of small-cell lung tumors. Chemokines do play a role in the recognition of tumorigenic cells (79), and chemokine receptors might play a role in lung cancer.

The more likely candidates for CCR influence are inflammatory diseases, such as rheumatoid arthritis, asthma, and inflammatory bowel disease (IBD). Preliminary studies have shown associations between the severity of rheumatoid arthritis and the *CCR5-Δ32* allele (39, 42). Higher levels of CCR5 expression on T cells have been observed in relapsing multiple sclerosis patients (5). The *CCR5-Δ32*

**TABLE 4** Diseases in the 3p region*

| Symbol | Name | Location |
|---|---|---|
| SCLC1 | Small-cell lung cancer | 3p23-21 |
| ARRVD5 | Arrhythmogenic right ventricular dysplasia | 3p23 |
| | Larssen syndrome 1 (developmental, skeletal) | 3p21.1-p14.1 |
| AS1 | Familial infantile encephalopathy | 3p21 |
| NCIE | Nonbullous congenital ichthyosiform erythroderma | 3p21 |
| DFNB6 | Autosomal recessive deafness | 3p21-p14 |

*Data from OMIM database (http://www.ncbi.nlm.nih.gov/OMIM).

allele is not associated with multiple sclerosis (MS) susceptibility (9) but is associated with a lower risk of MS reoccurrence (120) and delayed age-of-onset (6). An association between the CCR5 deletion and asthma was reported (48) but was not confirmed in subsequent studies (96, 119a). The CCR5 deletion has also been linked to pulmonary sarcoidosis (109), and negative studies for IBD and polymyalgia rheumatica have been published (23a, 86, 113a, 118). The CCR5-Δ32 allele was not associated with insulin dependent diabetes, whereas the CCR2-64I allele did show a modest association (128).

Although *CCR5-Δ32* homozygotes are otherwise healthy, an analysis of a cohort of homozygous hemophiliacs found slight differences in total lymphocyte counts and an increased incidence of hypertension (99). A comprehensive analysis of all of the chemokine receptor genes for variants, functional analysis of their products, and their association with disease will be required to fully elucidate the role of these genes in chronic human disease conditions.

## HLA, HIV-1, AND INFECTIOUS DISEASES

The human major histocompatibility complex (MHC) located on chromosome 6p21.31 contains the most polymorphic set of genes known in humans, the *HLA* class I and II loci (8, 71, 107). Products of these genes bind antigenic peptides and present them to T cells, initiating an immune response and clearance of the foreign material. The extraordinary polymorphism and balanced distribution of allelic frequencies characterizing the *HLA* loci are believed to be maintained through selective forces such as infectious disease morbidity (59, 108). Selective maintenance of MHC diversity is strongly implicated by evolutionary and theoretical data, including analyses of *HLA* allele frequency distribution in populations (49, 50), the high incidence of nonsynonymous base substitutions in regions of the genes that encode the peptide-binding domains (57, 58), persistence for millions of years of many transpecific polymorphisms (72, 129), and increased disease susceptibility in species with limited MHC diversity (11, 12, 33, 106, 139).

### Heterozygote Advantage

The hypothesis of overdominant selection (heterozygote advantage) for *HLA* proposes that individuals heterozygous at *HLA* loci are able to present a greater variety of antigenic peptides than are homozygotes, resulting in a more effective immune response to a wide diversity of pathogens (30, 144). Alternatively, homozygosity at these loci would limit the repertoire of interactions between HLA molecules and T lymphocytes. Theoretically, protective effects of heterozygosity would be most obvious at a population level against a pathogen that undergoes a fairly high rate of mutation because, on the average, a mutant that is capable of escaping immune surveillance would take longer to arise in heterozygotes than it would in homozygotes at these loci (100, 103, 110).

A recent genetic epidemiological study revealing overdominant selection at the *HLA* class I loci entailed an analysis of 498 HIV-1 positive seroconverters whose infection date was known within a six-month period (20). A strong association between *HLA* class I homozygosity (at one or more loci) and rapid progression to AIDS was observed in both Caucasian and African American study cohorts. The effect was particularly striking in individuals who were homozygous at two or three loci, a situation that is usually predictive of rapid progression on an individual basis. All three class I loci appeared to contribute to the association. Similar increased rates for AIDS progression were evident among homozygotes at *HLA-A* and *HLA-B* in 140 Dutch homosexual men and 202 Rwandan heterosexual women infected with HIV-1 (130). There was a stronger association with homozygosity at the *HLA-B* locus in the Amsterdam cohort and at the *HLA-A* locus in the Rwandan cohort, consistent with previous evidence that *HLA-A* and *HLA-B* genes are subject to varying degrees of gene flow and natural selection in human populations (112). The most parsimonious explanation for these data is that heterozygotes are able to present a broader range of HIV-1 peptides, thereby prolonging the time it takes for an escape mutant to arise. Similar epidemiological benefits for DR-DQ heterozygosity in clearing hepatitis B virus has been noted (134), suggesting that the heterozygote advantage model may also apply to the *HLA* class II loci. Thus, epidemiologic data supporting a role for infectious diseases in selection for heterozygosity at the *HLA* loci is beginning to accumulate.

## HLA Associations with Infectious Diseases

A search for the genetic effects of specific *HLA* class I and class II alleles on infectious diseases has been ongoing for several decades; and among the variety of associations observed (52), a few consistent findings have been identified. Although *HLA* loci are likely candidates to influence infectious disease pathogenesis, host genetic effects on most infectious diseases may well be multifactorial. The incomplete penetrance of *HLA*, along with the extreme polymorphism of these loci, necessitates the use of very large sample sizes in order to achieve adequate power for assessment. These issues have been the Achilles heel of many HLA and infectious disease studies, although limitations in patient clinical descriptions, failure to correct for multiple comparisons, and the low-resolution HLA typing methods have also added to the confusion (105).

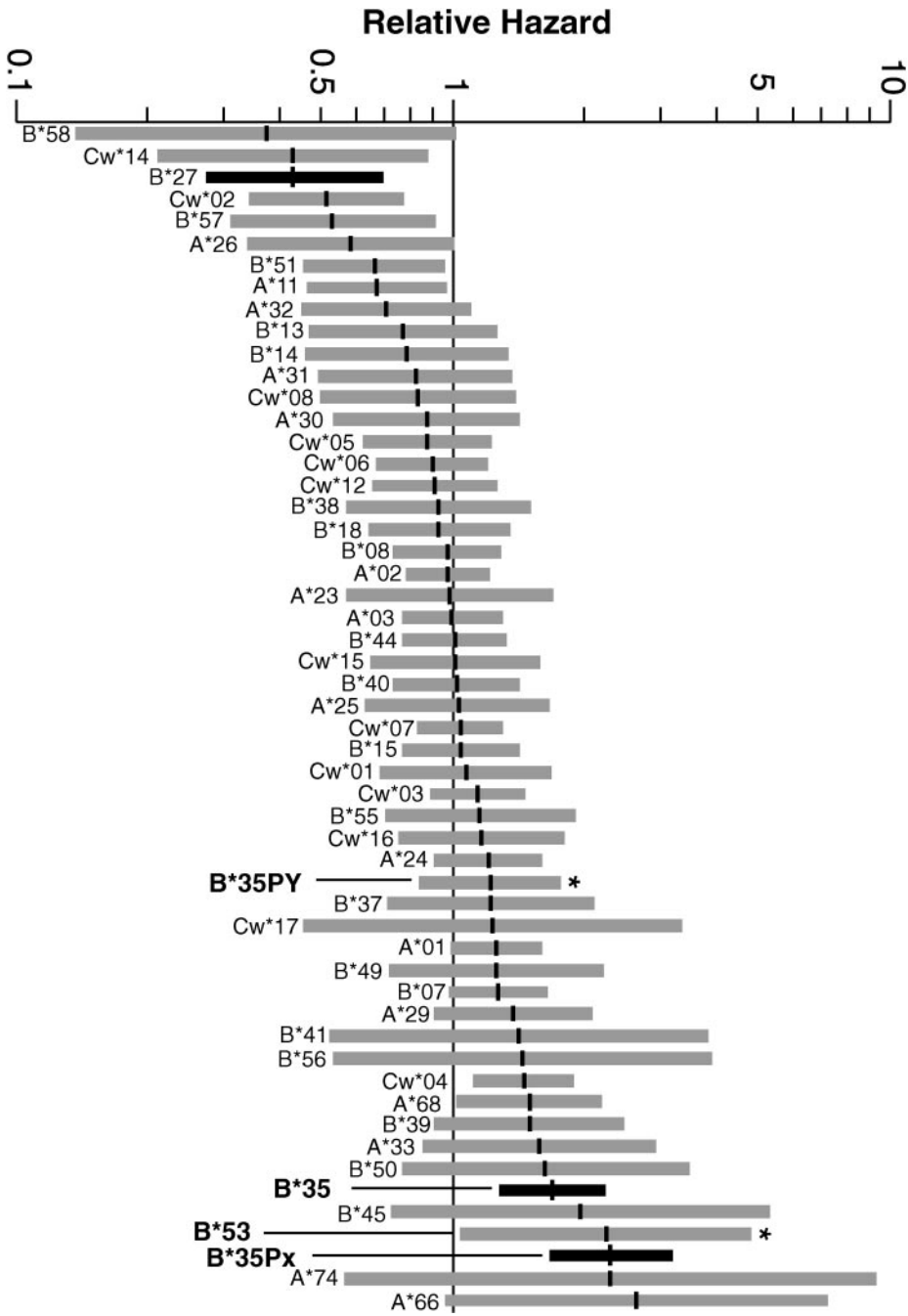
Some of the earliest work regarding *HLA* and infectious diseases involved leprosy and tuberculosis, where HLA-DR2 associations with both diseases have been reported (15, 16, 70, 125, 135, 138). The most convincing effect of HLA on a nonviral infectious disease was reported for B*5301 protection from severe malaria in Gambian children (53). A protective effect of DRB1*1302 was also observed. One of the hallmarks of this study was the large sample of individuals tested, a key component in discriminating true associations from those that are spurious. No B*5301 effect was observed in a similar study of severe malaria in Kenya (52), perhaps owing to genetic differences in the malaria parasite endemic in Kenya versus in The Gambia.

HLA associations with several viral infections, including hepatitis B virus (HBV), hepatitis C virus (HCV), human T-lymphotropic virus-I (HTLV-I), and HIV, have been reported (2, 7, 24, 55, 62, 95, 131–133). Like the protective effect against severe malaria, DRB1*1302 was associated with HBV clearance in Gambians (55, 133), and this is probably the most consistent finding among all studies of HLA effects on HBV persistence (131). The strongest HLA associations with HCV have also involved class II loci, where DQB1*0301 association with viral clearance has been replicated in a number of studies (2, 24, 95, 132), including a recent study of 200 HCV clearance and 374 matched persistently HCV-infected subjects (131a). A particularly strong association between DRB1*0101 and viral clearance in Caucasians was reported in the latter study, supporting the protective effect of this allele in a cohort of Irish women reported previously (7). DRB1*0101 was associated with disease susceptibility in carriers of HTLV-I, whereas HLA-A*02 strongly protected against HTLV-I-associated myelopathy (62). A growing number of reports convincingly indicate that development of cancer or neoplasia after infection with human papilloma virus (HPV) is also linked to specific HLA types (138a).

Over the past decade, HIV-1 disease has been the most widely studied infectious disease in terms of genetic interaction with HLA loci. Many reports describing a role for *HLA* alleles or haplotypes in HIV infection, rate of disease progression, and disease sequelae have been published (64). Among the reported effects of genetic polymorphisms on rate of progression to AIDS after HIV infection, susceptibility conferred by haplotypes composed of *Cw*04-B*35* has been the strongest and most significant (20, 61, 67, 73, 116) (Figure 4). We observed a strong effect of *B*35-Cw*04* on rapid disease progression in a study of 330 Caucasian and 144 African American seroconverters (20) and showed that the *B*35-Cw*04* haplotype has a codominant effect, in that homozygotes for this haplotype progress more rapidly than do heterozygotes and *B*35-Cw*04* heterozygotes progress more rapidly than individuals without this haplotype (20).

Recently, the effect of the *B*35-Cw*04* haplotype had been attributed to a subset of B*35 alleles, excluding the most common subtype, B*3501 (38) (Figure 4). A total of 592 Caucasians and 219 African Americans were typed using high-resolution techniques. B*35 subtypes were divided into two groups based on peptide binding specificity: (*a*) *B*35-PY* group, consisting primarily of *B*3501*, which recognizes epitopes with proline in position 2 and tyrosine in position 9; and (*b*) a more broadly reactive *B*35-Px* group (B*3502, 3503, 3504, and the closely

**Figure 4** Associations of HLA class 1 types with progression to AIDS. Relative hazard (RH) values were determined for each of the 54 *HLA* types used. Bars represent 95% confidence intervals (CI) for each RH. Class I types with CI that cross the line representing a RH of 1 are not significant, whereas all others are significant at $p < 0.05$. All *B*35* alleles combined are significantly associated with progression to AIDS, as is *B*53*. However, *B*35PY*, a subset of the *B*35* group of alleles, is not significant, whereas the *B*35Px* subgroup is strongly associated with AIDS progression.


Relative Hazard
0.1
0.5
1
5
10
B*58
Cw*14
B*27
Cw*02
B*57
A*26
B*51
A*11
A*32
B*13
B*14
A*31
Cw*08
A*30
Cw*05
Cw*06
Cw*12
B*38
B*18
B*08
A*02
A*23
A*03
B*44
Cw*15
B*40
A*25
Cw*07
B*15
Cw*01
Cw*03
B*55
Cw*16
A*24
B*35PY
*
B*37
Cw*17
A*01
B*49
B*07
A*29
B*41
B*56
Cw*04
A*68
B*39
A*33
B*50
B*35
B*45
B*53
*
B*35Px
A*74
A*66

related B*5301 allele) that binds epitopes with proline in position 2 but accepts several different amino acids, excluding tyrosine, at position 9. Survival analysis indicated that the accelerating influence of *B*35* on progression to AIDS was completely attributable to the *B35-Px* alleles, some of which differ from *B*35-PY* alleles by only one amino acid residue. This study leaves little doubt as to the inadequacy of B*35-Px in controlling HIV-1 disease, relative to other alleles, and affirms the importance of distinct epitope specificity of closely related class I molecules in immune defense against HIV-1.

Viral epitopes for both B*3501 and Cw*04 molecules have been identified (63, 115, 121, 136), but no viral epitopes have been reported for the other B35 subtypes (Los Alamos HIV Molecular Immunology Database, http://iv-web.lanl.gov/immunology). Given the association of the *B*35-Cw*04* alleles with progression to AIDS, it will be interesting to determine whether these alleles are inferior in some way in inducing a productive CTL response. Preliminary data suggest that the strength of the total CTL response cannot account for the rapid progression to AIDS in these individuals (X. Jin, unpublished data). Perhaps a qualitative distinction in CTL activity accounts for the difference in rate of progression to AIDS between individuals with *B*35-Cw*04*, relative to other haplotypes.

*HLA-B*27* was predicted to be protective against HIV-1 because it presents an HIV-1 peptide epitope that does not readily undergo mutation (43, 68a, 110). A recent analysis of 592 seroconverters in our AIDS cohorts indicated that, indeed, *B*27* has a protective effect on progression to AIDS ($RH = 0.43$, $p = 0.001$), although this was not significant after correction for multiple tests (38). A high frequency of *HLA-B*57* was recently reported in a group of 13 HIV-positive long-term nonprogressors (85%), relative to individuals with progressive disease (9.5%) (93), in agreement with previous epidemiologic reports (67, 68). This association with protection was supported by genetic data derived from the 592 seroconverters ($RH = 0.55$, $p = 0.04$), though this was not significant after correction (38). Many other associations between *HLA* haplotype and the rate of HIV disease progression have been reported, but most of these findings have been difficult to confirm in independent or combined cohorts.

Genetic data have also suggested a protective role for *HLA* in HIV infection and transmission. Concordance of HLA class I types between mother-infant pairs was associated with a stepwise increase in the risk of perinatal HIV-1 transmission with each additional concordant allele (82). The HLA A2/6802 supertype, a group of alleles sharing binding preference for some peptide epitopes (27, 123, 124), was also associated with decreased risk of perinatal HIV-1 infection (83). A subset of this group of alleles had previously been found in significantly higher frequency in a group of highly exposed HIV-1 negative female sex workers from Nairobi, relative to the HIV-1 positive individuals in the same cohort (84). DRB1*01 was also associated with resistance to infection in this group. These data provide further evidence supporting the importance of an immunologically mediated mechanism of resistance to HIV-1.

Thus, the advent of molecular typing techniques and organization of large, clinically well-defined cohorts have begun to shed light on the complex genetic

relationship between infectious diseases and genes residing within the MHC. The growing body of literature clarifying the genetic effects of HLA on infectious diseases is beginning to illustrate the opposing selection pressures on particular alleles, as one might hypothesize for B*5301, where positive and negative pressure owing to malaria and HIV-1 disease, respectively, appear to exist. The genetic data regarding HLA effects on HIV-1 pathogenesis complement the wealth of cellular data that have been reported, and together, they strongly emphasize the critical influence of this locus on controlling HIV disease.

## PERSPECTIVE

Both chemokine receptors and HLA proteins play critical roles in the immune response and provide an opportunity to understand the influence of pathogens on the evolution of the human genome. As described above, both gene clusters are large, contain many duplicated genes, and have interesting patterns of variation and LD. Much more work needs to be done on the CCR region, including completion of the sequence, identification of variation (SNPs), functional analysis of SNPs, and assessment of association to disease and LD. It is likely that much more evidence will be found for influence of the CCR genes on pathogen resistance and complex genetic disease.

Although a diverse and powerful immune response to foreign agents and damaged/mutated cells is desirable, it is becoming clear that this is a double-edged sword. The immune response against infected or damaged cells can lead to inflammation that over time can lead to medical complications. In addition, autoimmune reactions are important contributors to diseases such as lupus, arthritis, diabetes, etc. The development of vaccines and antibiotics against many of the most deadly human pathogens has resulted in a greatly decreased mortality rate and an increase in the average life span. At the same time, there is a greatly increasing frequency of many diseases with an inflammatory or autoimmune component (NIDDM, MS, etc.). The immune mechanisms that allowed our species to survive until the present may have contributed to a number of common complex diseases. Thus, a better understanding of the role of immune response genes, their evolution and variation, is likely to benefit the understanding of infectious disease as well as many other disorders.

**The *Annual Review of Genomics and Human Genetics* is online at http://genom.annualreviews.org**

## LITERATURE CITED

1. Adams RJ. 1995. Sickle cell disease and stroke. *J. Child Neurol.* 10:75–76
2. Alric L, Fort M, Izopet J, Vinel JP, Charlet JP, et al. 1997. Genes of the major histocompatibility complex class II influence the outcome of hepatitis C virus infection. *Gastroenterology* 113:1675–81

3. Angelicheva D, Boteva K, Jordanova A, Savov A, Kufardjieva A, et al. 1994. Cystic fibrosis patients from the Black Sea region: the 1677delTA mutation. *Hum. Mutat.* 3:353–57
4. Ansari-Lari MA, Liu XM, Metzker ML, Rut AR, Gibbs RA. 1997. The extent of genetic variation in the CCR5 gene. *Nat. Genet.* 16:221–22
5. Balashov KE, Rottman JB, Weiner HL, Hancock WW. 1999. CCR5(+) and CXCR3(+) T cells are increased in multiple sclerosis and their ligands MIP-1alpha and IP-10 are expressed in demyelinating brain lesions. *Proc. Natl. Acad. Sci. USA* 96:6873–78
6. Barcellos LF, Schito AM, Rimmler JB, Vittinghoff E, Shih A, et al. 2000. CC-chemokine receptor 5 polymorphism and age of onset in familial multiple sclerosis. Multiple Sclerosis Genetics Group. *Immunogenetics* 51:281–88
7. Barrett S, Ryan E, Crowe J. 1999. Association of the HLA-DRB1*01 allele with spontaneous viral clearance in an Irish cohort infected with hepatitis C virus via contaminated anti-D immunoglobulin. *J. Hepatol.* 30:979–83
8. Beck S, Trowsdale J. 2000. The human major histocompatibility complex: lessons from the DNA sequence. *Annu. Rev. Genomics Hum. Genet.* 1:117–37
9. Bennetts BH, Teutsch SM, Buhler MM, Heard RN, Stewart GJ. 1997. The CCR5 deletion mutation fails to protect against multiple sclerosis. *Hum. Immunol.* 58:52–59
10. Berger EA, Murphy PM, Farber JM. 1999. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annu. Rev. Immunol.* 17: 657–700
11. Black FL. 1992. Why did they die? *Science* 258:1739–40
12. Black FL, Hedrick PW. 1997. Strong balancing selection at HLA loci: evidence from segregation in South Amerindian families. *Proc. Natl. Acad. Sci. USA* 94: 12452–56
13. Blanpain C, Lee B, Tackoen M, Puffer B, Boom A, et al. 2000. Multiple nonfunctional alleles of CCR5 are frequent in various human populations. *Blood* 96: 1638–45
14. Boring L, Gosling J, Monteclaro FS, Lusis AJ, Tsou CL, Charo IF. 1996. Molecular cloning and functional expression of murine JE (monocyte chemoattractant protein 1) and murine macrophage inflammatory protein 1alpha receptors: evidence for two closely linked C-C chemokine receptors on chromosome 9. *J. Biol. Chem.* 271:7551–58
15. Bothamley GH, Beck JS, Schreuder GM, D'Amaro J, de Vries RR, et al. 1989. Association of tuberculosis and M. tuberculosis-specific antibody levels with HLA. *J. Infect. Dis.* 159:549–55
16. Brahmajothi V, Pitchappan RM, Kakkanaiah VN, Sashidhar M, Rajaram K, et al. 1991. Association of pulmonary tuberculosis and HLA in South India. *Tuber* 72:123–32
17. Bream JH, Young HA, Rice N, Martin MP, Carrington M, O'Brien SJ. 1999. *CCR5* promoter alleles and specific DNA binding factors. *Science* 284:223a
18. Cairns JS, D'Souza MP. 1998. Chemokines and HIV-1 second receptors: the therapeutic connection. *Nat. Med.* 4: 563–68
19. Carrington M, Kissner T, Gerrard B, Ivanov S, O'Brien SJ, Dean M. 1997. Novel alleles of the chemokine-receptor gene *CCR5*. *Am. J. Hum. Genet.* 61:1261–67
20. Carrington M, Nelson GW, Martin MP, Kissner T, Vlahov D, et al. 1999. HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science* 283:1748–52
21. Charo IF, Myers SJ, Herman A, Franci C, Connolly AJ, Coughlin SR. 1994. Molecular cloning and functional expression of two monocyte chemoattractant protein

1 receptors reveals alternative splicing of the carboxyl-terminal tails. *Proc. Natl. Acad. Sci. USA* 91:2752–56

22. Chaudhuri A, Polyakova J, Zbrzezna V, Pogo AO. 1995. The coding sequence of Duffy blood group gene in humans and simians: restriction fragment length polymorphism, antibody and malarial parasite specificities, and expression in nonerythroid tissues in Duffy-negative individuals. *Blood* 85:615–21
23. Choe H, Farzan M, Sun Y, Sullivan N, Rollins B, et al. 1996. The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* 85:1135–48

23a. Craggs A, Welfare M, Donaldson PT, Mansfield JC. 2001. The CC chemokine receptor 5-delta32 mutation is not associated with inflammatory bowel disease (IBD) in NE England. *Genes Immun.* 2:114–16

24. Cramp ME, Carucci P, Underhill J, Naoumov NV, Williams R, Donaldson PT. 1998. Association between HLA class II genotype and spontaneous clearance of hepatitis C viraemia. *J. Hepatol.* 29:207–13
25. Cuthbert AW, Halstead J, Ratcliff R, Colledge WH, Evans MJ. 1995. The genetic advantage hypothesis in cystic fibrosis heterozygotes: a murine study. *J. Physiol.* (London) 482:449–54
26. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, et al. 1996. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science* 273:1856–62
27. del Guercio MF, Sidney J, Hermanson G, Perez C, Grey HM, et al. 1995. Binding of a peptide antigen to multiple HLA alleles allows definition of an A2-like supertype. *J. Immunol.* 154:685–93
28. Deng HK, Unutmaz D, Kewal-Ramani VN, Littman DR. 1997. Expression cloning of new receptors used by simian and human immunodeficiency viruses. *Nature* 388:296–300
29. Diamond J. 1991. Curse and blessing of the ghetto. *Discover* 12:60–65
30. Doherty PC, Zinkernagel RM. 1975. Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature* 256:50–52
31. Doranz BJ, Rucker J, Yi Y, Smyth RJ, Samson M, et al. 1996. A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell* 85:1149–58
32. Estivill X, Bancells C, Ramos C. 1997. Geographic distribution and regional origin of 272 cystic fibrosis mutations in European populations. The Biomed CF Mutation Analysis Consortium. *Hum. Mutat.* 10:135–54
33. Evans DT, Piekarczyk MS, Allen TM, Boyson JE, Yeager M, et al. 1997. Immunodominance of a single CTL epitope in a primate species with limited MHC class I polymorphism. *J. Immunol.* 159:1374–82

33a. Faure S, Meyer L, Costagliola D, Vaneensberghe C, Genin E, et al. 2000. Rapid progression to AIDS in HIV+ individuals with a structural variant of the chemokine receptor CX3CR1. *Science* 287:2274–77

34. Flint J, Hill AV, Bowden DK, Oppenheimer SJ, Sill PR, et al. 1986. High frequencies of alpha-thalassaemia are the result of natural selection by malaria. *Nature* 321:744–50
35. Friedman MJ, Trager W. 1981. The biochemistry of resistance to malaria. *Sci. Am.* 244:154–55, 158–64
36. Gabriel SE, Clarke LL, Boucher RC, Stutts MJ. 1993. CFTR and outward rectifying chloride channels are distinct proteins with a regulatory relationship. *Nature* 363:263–68

37. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, et al. 1999. Origin of HIV-1 in the chimpanzee Pan troglodytes. *Nature* 397:436–41
38. Gao X, Nelson GW, Karacki P, Martin MP, Phair J, et al. 2001. Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. *N. Engl. J. Med.* 344:1668–75
39. Garred P, Madsen HO, Petersen J, Marquart H, Hansen TM, et al. 1998. CC chemokine receptor 5 polymorphism in rheumatoid arthritis. *J. Rheumatol.* 25: 1462–65
40. Garzino-Demo A, DeVico AL, Conant KE, Gallo RC. 2000. The role of chemokines in human immunodeficiency virus infection. *Immunol. Rev.* 177:79–87
41. Girardet A, McPeek MS, Leeflang EP, Munier F, Arnheim N, et al. 2000. Meiotic segregation analysis of RB1 alleles in retinoblastoma pedigrees by use of single-sperm typing. *Am. J. Hum. Genet.* 66:167–75
42. Gomez-Reino JJ, Pablos JL, Carreira PE, Santiago B, Serrano L, et al. 1999. Association of rheumatoid arthritis with a functional chemokine receptor, CCR5. *Arthritis Rheum.* 42:989–92
43. Goulder PJ, Phillips RE, Colbert RA, McAdam S, Ogg G, et al. 1997. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nat. Med.* 3: 212–17
44. Grebner EE, Tomczak J. 1991. Distribution of three alpha-chain beta-hexosaminidase A mutations among Tay-Sachs carriers. *Am. J. Hum. Genet.* 48:604–7
45. Guignard F, Combadiere C, Tiffany HL, Murphy PM. 1998. Gene organization and promoter function for CC chemokine receptor 5 (CCR5). *J. Immunol.* 160:985–92
46. Hadley TJ, Peiper SC. 1997. From malaria to chemokine receptor: the emerging physiologic role of the Duffy blood group antigen. *Blood* 89:3077–91
47. Haldane JBS. 1949. Disease and evolution. *Ric. Sci. Suppl.* 19:2–11
48. Hall IP, Wheatley A, Christie G, McDougall C, Hubbard R, Helms PJ. 1999. Association of CCR5 delta32 with reduced risk of asthma. *Lancet* 354:1264–65
49. Hedrick PW, Thomson G. 1983. Evidence for balancing selection at HLA. *Genetics* 104:449–56
50. Hedrick PW, Whittam TS, Parham P. 1991. Heterozygosity at individual amino acid sites: extremely high levels for HLA-A and -B genes. *Proc. Natl. Acad. Sci. USA* 88:5897–901
51. Herrick JB. 1910. Peculiar elongated and sickle-shaped red blood corpuscles in a case of severe anemia. *Arch. Intern. Med.* 6:517–21
52. Hill AV. 1998. The immunogenetics of human infectious diseases. *Annu. Rev. Immunol.* 16:593–617
53. Hill AV, Allsopp CE, Kwiatkowski D, Anstey NM, Twumasi P, et al. 1991. Common West African HLA antigens are associated with protection from severe malaria. *Nature* 352:595–600
54. Hogenauer C, Santa Ana CA, Porter JL, Millard M, Gelfand A, et al. 2000. Active intestinal chloride secretion in human carriers of cystic fibrosis mutations: an evaluation of the hypothesis that heterozygotes have subnormal active intestinal chloride secretion. *Am. J. Hum. Genet.* 67:1422–27
55. Hohler T, Gerken G, Notghi A, Lubjuhn R, Taheri H, et al. 1997. HLA-DRB1*1301 and *1302 protect against chronic hepatitis B. *J. Hepatol.* 26:503–7
56. Howard OM, Shirakawa AK, Turpin JA, Maynard A, Tobin GJ, et al. 1999. Naturally occurring CCR5 extracellular and transmembrane domain variants affect HIV-1 Co-receptor and ligand binding function. *J. Biol. Chem.* 274:16228–34

57. Hughes AL, Nei M. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* 335:167–70
58. Hughes AL, Nei M. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc. Natl. Acad. Sci. USA* 86:958–62
59. Hughes AL, Yeager M. 1998. Natural selection at major histocompatibility complex loci of vertebrates. *Annu. Rev. Genet.* 32:415–35
60. Ingram VM. 1959. A specific elemental difference between the globins of normal and sickle cell anemia hemoglobins. *Nature* 178:792–94
61. Itescu S, Mathur-Wagh U, Skovron ML, Brancato LJ, Marmor M, et al. 1992. HLA-B35 is associated with accelerated progression to AIDS. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* 5:37–45
62. Jeffery KJ, Usuku K, Hall SE, Matsumoto W, Taylor GP, et al. 1999. HLA alleles determine human T-lymphotropic virus-I (HTLV-I) proviral load and the risk of HTLV-I–associated myelopathy. *Proc. Natl. Acad. Sci. USA* 96:3848–53
63. Johnson RP, Trocha A, Buchanan TM, Walker BD. 1993. Recognition of a highly conserved region of human immunodeficiency virus type 1 gp120 by an HLA-Cw4–restricted cytotoxic T-lymphocyte clone. *J. Virol.* 67:438–45
64. Just JJ. 1995. Genetic predisposition to HIV-1 infection and acquired immune deficiency virus syndrome: a review of the literature examining associations with HLA [corrected]. *Hum. Immunol.* 44:156–69
65. Kalman YM, Kerem E, Darvasi A, DeMarchi J, Kerem B. 1994. Difference in frequencies of the cystic fibrosis alleles, delta F508 and W1282X, between carriers and patients. *Eur. J. Hum. Genet.* 2:77–82
66. Kan YW, Dozy AM. 1978. Antenatal diagnosis of sickle-cell anaemia by DNA analysis of amniotic-fluid cells. *Lancet* 2:910–12
67. Kaslow RA, Carrington M, Apple R, Park L, Munoz A, et al. 1996. Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. *Nat. Med.* 2:405–11
67a. Kato H, Tsuchiya N, Izumi S, Miyamasu M, Nakajima T, et al. 1999. New variations of human CC-chemokine receptors CCR3 and CCR4. *Genes Immun.* 1:97–104
68. Keet IP, Klein MR, Just JJ, Kaslow RA. 1996. The role of host genetics in the natural history of HIV-1 infection: the needles in the haystack. *AIDS* 10:S59–67
68a. Kelleher AD, Long C, Holmes EC, Allen RL, Wilson J, et al. 2001. Clustered mutations in HIV-1 gag are consistently required for escape from HLA-B27-restricted cytotoxic T lymphocyte responses. *J. Exp. Med.* 193:375–86
69. Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, et al. 1989. Identification of the cystic fibrosis gene: genetic analysis. *Science* 245:1073–80
70. Khomenko AG, Litvinov VI, Chukanova VP, Pospelov LE. 1990. Tuberculosis in patients with various HLA phenotypes. *Tubercle* 71:187–92
71. Klein J. 1986. *Natural History of the Major Histocompatibility Complex*. New York: Wiley. 775 pp.
72. Klein J. 1987. Origin of major histocompatibility complex polymorphism: the *trans*-species hypothesis. *Hum. Immunol.* 19:155–62
73. Klein MR, Keet IP, D'Amaro J, Bende RJ, Hekman A, et al. 1994. Associations between HLA frequencies and pathogenic features of human immunodeficiency virus type 1 infection in seroconverters from the Amsterdam cohort of homosexual men. *J. Infect. Dis.* 169:1244–49

73a. Klitz W, Brautbar C, Schito AM, Barcellos LF, Oksenberg JR. 2001. Evolution of the *CCR5-?32* mutation based on haplotype variation in Jewish and Northern European population samples. *Hum. Immunol.* 62:530–38
74. Kostrikis LG, Huang Y, Moore JP, Wolinsky SM, Zhang L, et al. 1998. A chemokine receptor CCR2 allele delays HIV-1 disease progression and is associated with a CCR5 promoter mutation. *Nat. Med.* 4:350–53
75. Liao F, Alkhatib G, Peden KW, Sharma G, Berger EA, Farber JM. 1997. STRL33, a novel chemokine receptor-like protein, functions as a fusion cofactor for both macrophage-tropic and T cell line-tropic HIV-1. *J. Exp. Med.* 185:2015–23
76. Libert F, Cochaux P, Beckman G, Samson M, Aksenova M, et al. 1998. The delta-ccr5 mutation conferring protection against HIV-1 in Caucasian populations has a single and recent origin in Northeastern Europe. *Hum. Mol. Genet.* 7:399–406
77. Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, et al. 1996. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply exposed individuals to HIV-1 infection. *Cell* 86:367–77
78. Liu R, Zhao X, Gurney TA, Landau NR. 1998. Functional analysis of the proximal CCR5 promoter. *AIDS* 14:1509–19
79. Locati M, Murphy PM. 1999. Chemokines and chemokine receptors: biology and clinical relevance in inflammation and AIDS. *Annu. Rev. Med.* 50:425–40
80. Lucotte G, Mercier G. 1998. Distribution of the CCR5 gene 32-bp deletion in Europe. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* 19:174–77
81. Lucotte G, Mercier G, Smets P. 1999. Elevated frequencies of the mutant allele delta32 at the CCR5 gene in Ashkenazic Jews. *AIDS* 15:1–2
82. MacDonald KS, Embree J, Njenga S, Nagelkerke NJ, Ngatia I, et al. 1998. Mother-child class I HLA concordance increases perinatal human immunodeficiency virus type 1 transmission. *J. Infect. Dis.* 177:551–56
83. MacDonald KS, Embree JE, Nagelkerke NJ, Castillo J, Ramhadin S, et al. 2001. The HLA A2/6802 supertype is associated with reduced risk of perinatal human immunodeficiency virus type 1 transmission. *J. Infect. Dis.* 183:503–6
84. MacDonald KS, Fowke KR, Kimani J, Dunand VA, Nagelkerke NJ, et al. 2000. Influence of HLA supertypes on susceptibility and resistance to human immunodeficiency virus type 1 infection. *J. Infect. Dis.* 181:1581–89
85. Maho A, Bensimon A, Vassart G, Parmentier M. 1999. Mapping of the CCXCR1, CX3CR1, CCBP2, and CCR9 genes to the CCR cluster within the 3p21.3 region of the human genome. *Cytogenet. Cell Genet.* 87:265–68
86. Martin K, Heinzlmann M, Borchers R, Mack M, Loeschke K, Folwaczny C. 2001. Delta32 mutation of the chemokine-receptor 5 gene in inflammatory bowel disease. *Clin. Immunol.* 98:18–22
87. Martin MP, Dean M, Smith MW, Winkler C, Gerrard B, et al. 1998. Genetic acceleration of AIDS progression by a promoter variant of CCR5. *Science* 282:1907–11
88. Martinson JJ, Chapman NH, Rees DC, Liu YT, Clegg JB. 1997. Global distribution of the CCR5 gene 32–basepair deletion. *Nat. Genet.* 16:100–3
89. Martinson JJ, Hong L, Karanicolas R, Moore JP, Kostrikis LG. 2000. Global distribution of the CCR2-64I/CCR5-59653T HIV-1 disease-protective haplotype. *AIDS* 14:483–89
90. McDermott DH, Beecroft MJ, Kleeberger CA, Al-Sharif FM, Ollier WE, et al. 2000. Chemokine RANTES promoter polymorphism affects risk of both HIV infection and disease progression in the

Multicenter AIDS Cohort Study. *AIDS* 14:2671–78
91. McDermott DH, Zimmerman PA, Guignard F, Kleeberger CA, Leitman SF, Murphy PM. 1998. CCR5 promoter polymorphism and HIV-1 disease progression. Multicenter AIDS Cohort Study (MACS). *Lancet* 352:866–70
92. Mellado M, Rodriguez-Frade JM, Vila-Coro AJ, de Ana AM, Martinez AC. 1999. Chemokine control of HIV-1 infection. *Nature* 400:723–24
93. Migueles SA, Sabbaghian MS, Shupert WL, Bettinotti MP, Marincola FM, et al. 2000. HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long-term nonprogressors. *Proc. Natl. Acad. Sci. USA* 97:2709–14
94. Miller LH, Mason SJ, Clyde DF, McGinniss MH. 1976. The resistance factor to *Plasmodium vivax* in blacks. The Duffy–blood-group genotype, FyFy. *N. Engl. J. Med.* 295:302–4
95. Minton EJ, Smillie D, Neal KR, Irving WL, Underwood JC, James V. 1998. Association between MHC class II alleles and clearance of circulating hepatitis C virus. Members of the Trent Hepatitis C Virus Study Group. *J. Infect. Dis.* 178:39–44
96. Mitchell TJ, Walley AJ, Pease JE, Venables PJ, Wiltshire S, et al. 2000. Delta 32 deletion of CCR5 gene and association with asthma or atopy. *Lancet* 356:1491–92
97. Mummidi S, Ahuja SS, Gonzalez E, Anderson SA, Santiago EN, et al. 1998. Genealogy of the CCR5 locus and chemokine system gene variants associated with altered rates of HIV-1 disease progression. *Nat. Med.* 4:786–93
98. Mummidi S, Ahuja SS, McDaniel BL, Ahuja SK. 1997. The human CC chemokine receptor 5 (CCR5) gene. Multiple transcripts with 5′ end heterogeneity, dual promoter usage, and evidence for polymorphisms within the regulatory regions and noncoding exons. *J. Biol. Chem.* 272:30662–71
99. Nguyen GT, Carrington M, Beeler JA, Dean M, Aledort LM, et al. 1999. Phenotypic expressions of CCR5-delta32/delta32 homozygosity. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* 22:75–82
100. Nowak MA, Anderson RM, McLean AR, Wolfs TF, Goudsmit J, May RM. 1991. Antigenic diversity thresholds and the development of AIDS. *Science* 254:963–69
101. Oberlin E, Amara A, Bachelerie F, Bessia C, Virelizier JL, et al. 1996. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell–line-adapted HIV-1. *Nature* 382:833–35
102. O'Brien SJ. 1991. Ghetto legacy. *Curr. Biol.* 4:209–11
103. O'Brien SJ, Mann, DL. 1992. Adoptive chaos and AIDS. *Curr. Biol.* 2:203–5
104. O'Brien SJ, Moore JP. 2000. The effect of genetic variation in chemokines and their receptors on HIV transmission and progression to AIDS. *Immunol. Rev.* 177:99–111
105. O'Brien SJ, Nelson GW, Winkler CA, Smith MW. 2000. Polygenic and multifactorial disease gene association in man: lessons from AIDS. *Annu. Rev. Genet.* 34:563–91
106. O'Brien SJ, Roelke ME, Marker L, Newman A, Winkler CA, et al. 1985. Genetic basis for species vulnerability in the cheetah. *Science* 227:1428–34
107. Parham P, Adams EJ, Arnett KL. 1995. The origins of HLA-A,B,C polymorphism. *Immunol. Rev.* 143:141–80
108. Parham P, Ohta T. 1996. Population biology of antigen presentation by MHC class I molecules. *Science* 272:67–74
109. Petrek M, Drabek J, Kolek V, Zlamal J, Welsh KI, et al. 2000. CC chemokine receptor gene polymorphisms in Czech patients with pulmonary sarcoidosis. *Am. J. Respir. Crit. Care Med.* 162:1000–3

110. Phillips RE, Rowland-Jones S, Nixon DF, Gotch FM, Edwards JP, et al. 1991. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature* 354:453–59
111. Pier GB, Grout M, Zaidi T, Meluleni G, Mueschenborn SS, et al. 1998. *Salmonella typhi* uses CFTR to enter intestinal epithelial cells. *Nature* 393:79–82
112. Potts WK, Slev PR. 1995. Pathogen-based models favoring MHC genetic diversity. *Immunol. Rev.* 143:181–97
113. Quillent C, Oberlin E, Braun J, Rousset D, Gonzalez-Canali G, et al. 1998. HIV-1–resistance phenotype conferred by combination of two separate inherited mutations of CCR5 gene. *Lancet* 351:14–18
113a. Rector A, Vermeire S, Thoelen I, Keyaerts E, Struyf R, et al. 2001. Analysis of the CC chemokine receptor 5 (CCR5) delta-32 polymorphism in inflammatory bowel disease. *Hum. Genet.* 108:190–93
114. Rollins BJ. 1997. Chemokines. *Blood* 90:909–28
115. Rowland-Jones S, Sutton J, Ariyoshi K, Dong T, Gotch F, et al. 1995. HIV-specific cytotoxic T-cells in HIV-exposed but uninfected Gambian women. *Nat. Med.* 1:59–64
116. Sahmoud T, Laurian Y, Gazengel C, Sultan Y, Gautreau C, Costagliola D. 1993. Progression to AIDS in French hemophiliacs: association with HLA-B35. *AIDS* 7:497–500
117. Salemi M, Strimmer K, Hall WW, Duffy M, Delaporte E, et al. 2001. Dating the common ancestor of SIVcpz and HIV-1 group M and the origin of HIV-1 subtypes using a new method to uncover clock-like molecular evolution. *FASEB J.* 15:276–78
118. Salvarani C, Boiardi L, Timms JM, Silvestri T, Ranzi A, et al. 2000. Absence of the association with CC chemokine receptor 5 polymorphism in polymyalgia rheumatica. *Clin. Exp. Rheumatol.* 18: 591–95
119. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, et al. 1996. Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 382:722–25
119a. Sanford AJ, Zhu S, Bai TR, Fitzgerald JM, Pare PD. 2001. The role of the CC chemokine receptor-5 Delta32 polymorphism in asthma and in the production of regulated-on-activation normal T cells expressed and secreted. *J. Allergy Clin. Immunol.* 108:69–73
120. Sellebjerg F, Madsen HO, Jensen CV, Jensen J, Garred P. 2000. CCR5 delta32, matrix metalloproteinase-9 and disease activity in multiple sclerosis. *J. Neuroimmunol.* 102:98–106
121. Shiga H, Shioda T, Tomiyama H, Takamiya Y, Oka S, et al. 1996. Identification of multiple HIV-1 cytotoxic T-cell epitopes presented by human leukocyte antigen B35 molecules. *AIDS* 10:1075–83
122. Shin HD, Winkler C, Stephens JC, Bream J, Young H, et al. 2000. Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of IL10. *Proc. Natl. Acad. Sci. USA* 97:14467–72
123. Sidney J, Grey HM, Kubo RT, Sette A. 1996. Practical, biochemical and evolutionary implications of the discovery of HLA class I supermotifs. *Immunol. Today* 17:261–66
124. Sidney J, Grey HM, Southwood S, Celis E, Wentworth PA, et al. 1996. Definition of an HLA-A3–like supermotif demonstrates the overlapping peptide-binding repertoires of common HLA molecules. *Hum. Immunol.* 45:79–93
125. Singh SP, Mehra NK, Dingley HB, Pande JN, Vaidya MC. 1983. Human leukocyte antigen (HLA)-linked control of susceptibility to pulmonary tuberculosis and association with HLA-DR types. *J. Infect. Dis.* 148:676–81

126. Smith MW, Dean M, Carrington M, Winkler C, Huttley GA, et al. 1997. Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science* 277:959–65
127. Stephens JC, Reich DE, Goldstein DB, Shin HD, Smith MW, et al. 1998. Dating the origin of the CCR5-Delta32 AIDS-resistance allele by the coalescence of haplotypes. *Am. J. Hum. Genet.* 62:1507–15
128. Szalai C, Csaszar A, Czinner A, Szabo T, Panczel P, et al. 1999. Chemokine receptor CCR2 and CCR5 polymorphisms in children with insulin-dependent diabetes mellitus. *Pediatr. Res.* 46:82–84
129. Takahata N, Nei M. 1990. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* 124:967–78
130. Tang J, Costello C, Keet IP, Rivers C, Leblanc S, et al. 1999. HLA class I homozygosity accelerates disease progression in human immunodeficiency virus type 1 infection. *AIDS* 15:317–24
131. Thio CL, Thomas DL, Carrington M. 2000. Chronic viral hepatitis and the human genome. *Hepatology* 31:819–27
131a. Thio CL, Thomas DL, Goedert JJ, Vlahov D, Nelson KE, et al. 2001. Racial differences in HLA class II associations with hepatitus C virus outcomes. *J. Infect. Dis.* 184:16–21
132. Thursz M, Yallop R, Goldin R, Trepo C, Thomas HC. 1999. Influence of MHC class II genotype on outcome of infection with hepatitis C virus. Hepatitis C European Network for Cooperative Research. *Lancet* 354:2119–24
133. Thursz MR, Kwiatkowski D, Allsopp CE, Greenwood BM, Thomas HC, Hill AV. 1995. Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia. *N. Engl. J. Med.* 332:1065–69
134. Thursz MR, Thomas HC, Greenwood BM, Hill AV. 1997. Heterozygote advantage for HLA class-II type in hepatitis B virus infection. *Nat. Genet.* 17:11–12
135. Todd JR, West BC, McDonald JC. 1990. Human leukocyte antigen and leprosy: study in northern Louisiana and review. *Rev. Infect. Dis.* 12:63–74
136. Tomiyama H, Miwa K, Shiga H, Moore YI, Oka S, et al. 1997. Evidence of presentation of multiple HIV-1 cytotoxic T lymphocyte epitopes by HLA-B*3501 molecules that are associated with the accelerated progression of AIDS. *J. Immunol.* 158:5026–34
137. Tsui LC. 1995. The cystic fibrosis transmembrane conductance regulator gene. *Am. J. Respir. Crit. Care Med.* 151:S47–53
138. van Eden W, de Vries RR, Mehra NK, Vaidya MC, D'Amaro J, van Rood JJ. 1980. HLA segregation of tuberculoid leprosy: confirmation of the DR2 marker. *J. Infect. Dis.* 141:693–701
138a. Wang SS, Wheeler CM, Hildesheim A, Schiffman M, Herrero R, et al. 2001. Human leukocyte antigen class I and II alleles and risk of cervical neoplasia: results from a population-based study in Costa Rica. *J. Infect. Dis.* 184:1310–14
139. Watkins DI, Hodi FS, Letvin NL. 1988. A primate species with limited major histocompatibility complex class I polymorphism. *Proc. Natl. Acad. Sci. USA* 85:7714–18
140. Winkler C, Modi W, Smith MW, Nelson GW, Wu X, et al. 1998. Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. ALIVE Study, Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort. *Science* 279:389–93
141. Wright S. 1969. *Evolution and the*

*Genetics of Populations*. Chicago: Univ. Chicago Press

142. Zhang YW, Ryder OA, Zhang YP. 1999. Sequence evolution of the CCR5 chemokine receptor gene in primates. *Mol. Biol. Evol.* 16:1145–54
143. Zimmerman PA, Buckler-White A, Alkhatib G, Spalding T, Kubofcik J, et al. 1997. Inherited resistance to HIV-1 conferred by inactivating mutation in CC chemokine receptor 5: studies in populations with contrasting clinical phenotypes, defined racial background, and quantified risk. *Mol. Med.* 3:23–26
144. Zinkernagel RM. 1996. Immunology taught by viruses. *Science* 271:173–78

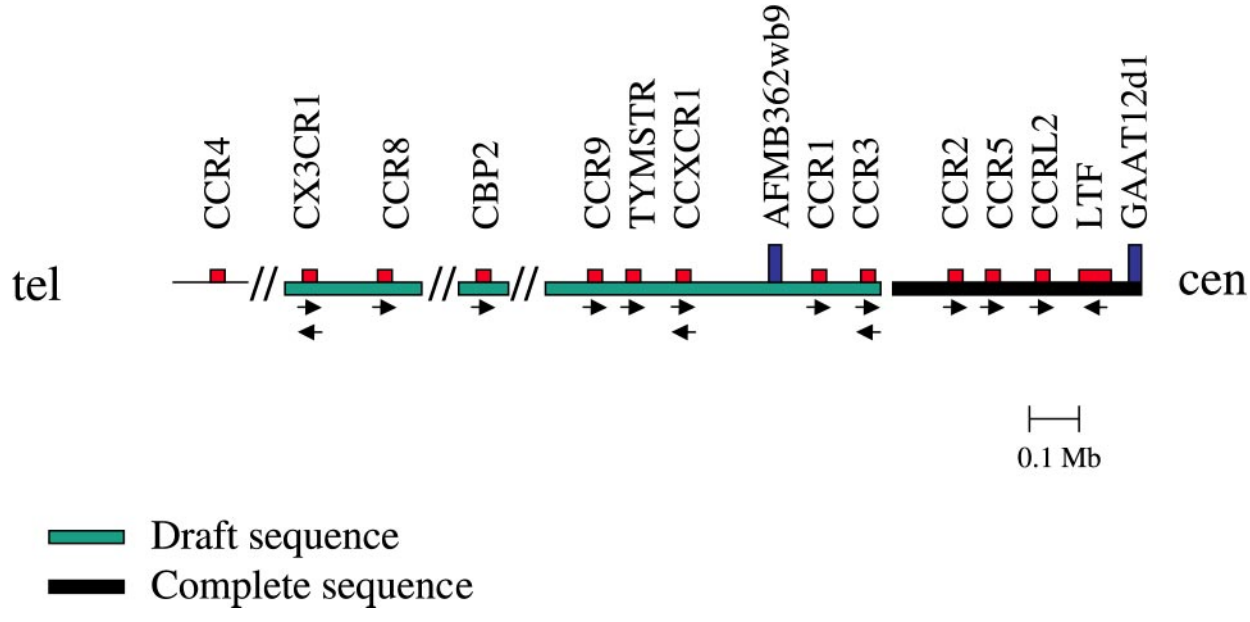



**Figure 2** Map of the CCR region. A map of the CCR region on chromosome 3p24-p21.31 region is shown. The order of each of the genes is shown, along with the orientation as determined by Human Genome Project Working Draft (http://genome.ucsc.edu/) or the NCBI Entrez viewer. For three genes (CX3CR1, CCXCR1, CCR3), the NCBI assembly predicts that they are on the opposite strand. The positioning of the genes on either complete sequence (black) or draft sequence (green) contigs is given, and gaps are shown between the contigs. The CCR4 gene is not located on any sequenced BAC to date. All of the genes are also present in the Celera draft sequence, except CCR4. AFMB362wb9, GAAT12d1, and microsatellites used in Stephens et al. (127). LFT, lactotransferrin gene; tel, telomere; cen, centromere.